Comparison of Interactive Voice Response SF-36 to Self-Administered SF-36 and Personal Interview via Telephone SF-36

David J. DeBrota, M.D., Edward W. Bradt, Carol M. Andrejasich, M.B.A.,
Mark Kosinski, M.A., and John E. Ware, Jr., Ph.D.
Eli Lilly and Company and The Health Institute, New England Medical Center

Health-related quality of life measures such as the SF-36 Health Status Survey have traditionally been accomplished with self-administered paper questionnaires and by interview administration. Increasingly affordable and easy-to-use computer technology offers new methods for collection of information directly from patients, with the potential to improve speed and quality while reducing the need for staff to do interviewing or data entry. For this study, an interactive voice response (IVR) system was developed to administer the SF-36 to patients over the telephone. The IVR system received inbound calls and made outbound calls, employed recorded voice prompts, and collected responses from subjects' presses on their touch-tone telephone keypads. The IVR method was compared with self-administration (SA) via optical mark recognition-ready paper forms returned by mail and with personal interview over the telephone (PIT) in a full crossover design. Subjects were to twice within 2-4 weeks have the SF-36 administered to them in their homes or workplaces, each time by a randomly-selected method (IVR, SA, or PIT). Subjects did not look at paper copies of the SF-36 questions during IVR or PIT administrations, and did not have their responses chosen in advance of telephone contact. Subjects were notified of administration methods to which they were randomized by means of letters mailed to them immediately prior to each of their administration date windows (each window was approx. 7-14 days long, and subjects could control when administration took place within it). Method preference and administration time information was collected, along with demographic and general health information. In order to maximize compliance with administrations, SA non-responders were mailed a second questionnaire if the first was not returned, and multiple telephone calls were made to PIT and IVR non-responders if necessary.

795 adults visiting outpatient medical clinics and an employee health clinic during the summer of 1995 agreed to participate. Mean age was 48 years (range: 18 to 87 years), 61% were female, 85% were caucasian, and 95% had graduated high school, with 54% being college graduates. These subjects completed a total of 1183 SF-36 administrations, with 541 subjects completing requested SF-36 administrations at both timepoints (T0 and T1). Response rates (completed/attempted) and mean administration times for method-timepoints were: IVR-T0: 228/309 (74%), 17 min; IVR-T1: 190/244 (78%), 16 min; PIT-T0: 226/244 (93%), 7 min; PIT-T1: 181/198 (91%), 7 min; SA-T0: 188/242 (78%), 10 min; SA-T1: 170/194 (88%), 12 min (9 min if an outlier reporting a 387 min administration time is removed). 376 subjects experienced two different methods and 214 of these expressed a method preference: SA was preferred over IVR (66 to 16), PIT was preferred over IVR (54 to 23), and SA was preferred over PIT (46 to 9).

The mean interadministration separation (T1-T0) was intended to be 14 days. The mean observed T1-T0 in all 541 subjects who completed T1 was 19 days (range: 10 to 65 days), and there were statistically significant differences in T1-T0 among method sequences. These differences can be attributed partially to human errors by study personnel and technical difficulties, and partially to effects of the methods and subject compliance.

SF-36 data obtained by all 3 methods was of high quality, with very few missing responses and good internal consistency. SF-36 scale scores were influenced by administration method and by method sequence, and possibly by T1-T0 differences. In general, PIT administration produced higher scores (suggestive of better health-related quality of life) than either IVR or SA, independent of method sequence. T1 scores were increased in both the IVR-SA and the SA-IVR method sequences.

It is possible that with continued improvements in interactive voice response technology, IVR administration of the SF-36 will provide an experience for subjects which is closer to that of a personal interview over the telephone. At present, however, IVR administration results in a lower response rate, requires more of subjects' time, and is not preferred by subjects, by comparison with conventional, well-validated administration methods. Furthermore, SF-36 scale scores from IVR administration may require adjustment in order to be compared with SF-36 scale scores from other administration methods.

0195—4210/96/$5.00